

#19



RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/478,598

DATE: 11/04/2002  
TIME: 16:19:28

Input Set : N:\Crf3\RULE60\09478598.raw  
Output Set: N:\CRF4\11042002\I478598.raw

1 <110> APPLICANT: Rao, Gururaj A.  
2 Major Sleister, Heidi  
3 <120> TITLE OF INVENTION: Compositions and Methods for Altering Amino Acid  
4 Content of Proteins  
5 <130> FILE REFERENCE: 5718-16  
6 <140> CURRENT APPLICATION NUMBER: 09/478,598  
7 <141> CURRENT FILING DATE: 2000-01-06  
9 <150> PRIOR APPLICATION NUMBER: US/08/988,015  
10 <151> PRIOR FILING DATE: 1997-12-10  
12 <160> NUMBER OF SEQ ID NOS: 11  
13 <170> SOFTWARE: PatentIn Ver. 2.0  
15 <210> SEQ ID NO: 1  
16 <211> LENGTH: 218  
17 <212> TYPE: PRT  
18 <213> ORGANISM: Glycine max  
19 <400> SEQUENCE: 1  
20 Arg Ser Ser Glu Val Lys Cys Ala Ser Phe Arg Leu Ala Val Glu Ala  
21 1 5 10 15  
22 His Asn Ile Arg Ala Phe Lys Thr Ile Pro Glu Glu Cys Val Ser Pro  
23 20 25 30  
24 Thr Lys Asp Tyr Ile Asn Gly Glu Gln Phe Arg Ser Asp Ser Lys Thr  
25 35 40 45  
26 Val Asn Gln Gln Ala Phe Phe Tyr Ala Ser Glu Arg Glu Val His His  
27 50 55 60  
28 Asn Asp Ile Phe Ile Phe Gly Ile Asp Asn Thr Val Leu Ser Asn Ile  
29 65 70 75 80  
30 Pro Tyr Tyr Glu Lys His Gly Tyr Gly Val Glu Glu Phe Asn Glu Thr  
31 85 90 95  
32 Leu Tyr Asp Glu Trp Val Asn Lys Gly Asp Ala Pro Ala Leu Pro Glu  
33 100 105 110  
34 Thr Leu Lys Asn Tyr Asn Lys Leu Leu Ser Leu Gly Phe Lys Ile Val  
35 115 120 125  
36 Phe Leu Ser Gly Arg Tyr Leu Asp Lys Met Ala Val Thr Glu Ala Asn  
37 130 135 140  
38 Leu Lys Lys Ala Gly Phe His Thr Trp Glu Gln Leu Ile Leu Lys Asp  
39 145 150 155 160  
40 Pro His Leu Ile Thr Pro Asn Ala Leu Ser Tyr Lys Ser Ala Met Arg  
41 165 170 175  
42 Glu Asn Leu Leu Arg Gln Gly Tyr Arg Ile Val Gly Ile Ile Gly Asp  
43 180 185 190  
44 Gln Trp Ser Asp Leu Leu Gly Asp His Arg Gly Glu Ser Arg Thr Phe  
45 195 200 205  
46 Lys Leu Pro Asn Pro Met Tyr Tyr Ile Glu

# ENTERED

## RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/478,598

TIME: 16:19:28

Input Set : N:\Crf3\RULE60\09478598.raw

Output Set: N:\CRF4\11042002\I478598.raw

```

47          210          215
49 <210> SEQ ID NO: 2
50 <211> LENGTH: 218
51 <212> TYPE: PRT
52 <213> ORGANISM: Glycine max
53 <400> SEQUENCE: 2
54   Arg Thr Pro Glu Val Lys Cys Ala Ser Trp Arg Leu Ala Val Glu Ala
55     1          5          10          15
56   His Asn Ile Phe Gly Phe Glu Thr Ile Pro Glu Glu Cys Val Glu Ala
57           20          25          30
58   Thr Lys Glu Tyr Ile His Gly Glu Gln Tyr Arg Ser Asp Ser Lys Thr
59       35          40          45
60   Val Asn Gln Gln Ala Tyr Phe Tyr Ala Arg Asp Leu Glu Val His Pro
61       50          55          60
62   Lys Asp Thr Phe Val Phe Ser Ile Asp Asn Thr Val Leu Ser Asn Ile
63       65          70          75          80
64   Pro Tyr Tyr Lys Lys His Gly Tyr Gly Val Glu Lys Phe Asn Ser Thr
65           85          90          95
66   Leu Tyr Asp Glu Trp Val Asn Lys Gly Asn Ala Pro Ser Leu Pro Glu
67       100         105         110
68   Thr Leu Lys Asn Tyr Asn Lys Leu Val Ser Leu Gly Phe Lys Ile Ile
69       115         120         125
70   Phe Leu Ser Gly Arg Thr Leu Asp Lys Gln Ala Val Thr Glu Ala Asn
71       130         135         140
72   Leu Lys Lys Ala Gly Tyr His Thr Trp Glu Lys Leu Ile Leu Lys Asp
73       145         150         155         160
74   Pro Gln Pro Ser Thr Pro Asn Ala Val Ser Tyr Lys Thr Ala Ala Arg
75           165         170         175
76   Glu Lys Leu Ile Arg Gln Gly Tyr Asn Ile Val Gly Ile Ile Gly Asp
77       180         185         190
78   Gln Trp Ser Asp Leu Leu Gly Gly His Arg Gly Glu Ser Arg Thr Phe
79       195         200         205
80   Lys Leu Pro Asn Pro Leu Tyr Tyr Ile Gln
81       210         215
83 <210> SEQ ID NO: 3
84 <211> LENGTH: 214
85 <212> TYPE: PRT
86 <213> ORGANISM: Lycopersicon esculentum
87 <400> SEQUENCE: 3
88   Leu Lys Cys Thr Thr Trp Arg Phe Val Val Glu Thr Asn Asn Leu Ser
89     1          5          10          15
90   Pro Trp Lys Thr Ile Pro Glu Glu Cys Ala Asp Tyr Val Lys Glu Tyr
91           20          25          30
92   Met Val Gly Pro Gly Tyr Lys Met Glu Ile Asp Arg Val Ser Asp Glu
93       35          40          45
94   Ala Gly Glu Tyr Ala Lys Ser Val Asp Leu Gly Asp Asp Gly Arg Asp
95       50          55          60
96   Val Trp Ile Phe Asp Val Asp Glu Thr Leu Leu Ser Asn Leu Pro Tyr
97       65          70          75          80

```

## RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/478,598

TIME: 16:19:28

Input Set : N:\Crif3\RULE60\09478598.raw

Output Set: N:\CRF4\11042002\I478598.raw

```

98      Tyr Ser Asp His Arg Tyr Gly Leu Glu Val Phe Asp Asp Val Glu Phe
99              85              90              95
100     Asp Lys Trp Val Glu Asn Gly Thr Ala Pro Ala Leu Gly Ser Ser Leu
101              100              105              110
102     Lys Leu Tyr Gln Glu Val Leu Lys Leu Gly Phe Lys Val Phe Leu Leu
103              115              120              125
104     Thr Gly Arg Ser Glu Arg His Arg Ser Val Thr Val Glu Asn Leu Met
105              130              135              140
106     Asn Ala Gly Phe His Asp Trp His Lys Leu Ile Leu Arg Gly Ser Asp
107     145              150              155              160
108     His Gly Lys Thr Ala Thr Thr Tyr Lys Ser Glu Arg Arg Asn Ala Met
109              165              170              175
110     Val Glu Glu Gly Phe Arg Ile Val Gly Asn Ser Gly Asp Gln Trp Ser
111              180              185              190
112     Asp Leu Leu Gly Ser Ser Met Ser Tyr Arg Ser Phe Lys Leu Pro Asn
113              195              200              205
114     Pro Met Tyr Tyr Ile Leu
115              210
117 <210> SEQ ID NO: 4
118 <211> LENGTH: 217
119 <212> TYPE: PRT
120 <213> ORGANISM: Phaseolus vulgaris
121 <400> SEQUENCE: 4
122     Ser Asp Thr Glu Val Arg Cys Ala Ser Trp Arg Leu Ala Val Glu Ala
123     1              5              10              15
124     Gln Asn Ile Phe Gly Phe Glu Thr Ile Pro Gln Gln Cys Val Asp Ala
125              20              25              30
126     Thr Ala Asn Tyr Ile Glu Gly Gly Gln Tyr Arg Ser Asp Ser Lys Thr
127              35              40              45
128     Val Asn Gln Gln Ile Tyr Phe Phe Ala Arg Asp Arg His Val His Glu
129              50              55              60
130     Asn Asp Val Ile Leu Phe Asn Ile Asp Gly Thr Ala Leu Ser Asn Ile
131              65              70              75              80
132     Pro Tyr Tyr Ser Gln His Gly Tyr Gly Ser Glu Lys Phe Asp Ser Glu
133              85              90              95
134     Arg Tyr Asp Glu Glu Phe Val Asn Lys Gly Glu Ala Pro Ala Leu Pro
135              100              105              110
136     Glu Thr Leu Lys Asn Tyr Asn Lys Leu Val Ser Leu Gly Tyr Lys Ile
137              115              120              125
138     Ile Phe Leu Ser Gly Arg Leu Lys Asp Lys Arg Ala Val Thr Glu Ala
139              130              135              140
140     Asn Leu Lys Lys Ala Gly Tyr Asn Thr Trp Glu Lys Leu Ile Leu Lys
141     145              150              155              160
142     Asp Pro Ser Asn Ser Ala Glu Asn Val Val Tyr Lys Thr Ala Glu Arg
143              165              170              175
144     Ala Lys Leu Val Gln Glu Gly Tyr Arg Ile Val Gly Asn Ile Gly Asp
145              180              185              190
146     Gln Trp Asn Asp Leu Lys Gly Glu Asn Arg Ala Ile Arg Ser Phe Lys
147              195              200              205

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/478,598

DATE: 11/04/2002

TIME: 16:19:28

Input Set : N:\Crf3\RULE60\09478598.raw

Output Set: N:\CRF4\11042002\I478598.raw

```

148      Leu Pro Asn Pro Met Tyr Tyr Thr Lys
149          210                      215
151 <210> SEQ ID NO: 5
152 <211> LENGTH: 214
153 <212> TYPE: PRT
154 <213> ORGANISM: Arabidopsis thaliana
155 <400> SEQUENCE: 5
156      Pro Asn Cys Arg Ser Trp His Leu Gly Phe Glu Thr Ser Asn Met Ile
157          1                      5                      10                      15
158      Asn Phe Asp Thr Val Pro Ala Asn Cys Lys Ala Tyr Val Glu Asp Tyr
159          20                      25                      30
160      Leu Ile Thr Ser Lys Gln Tyr Gln Tyr Asp Ser Lys Thr Val Asn Lys
161          35                      40                      45
162      Glu Ala Tyr Phe Tyr Ala Lys Gly Leu Ala Leu Lys Asn Asp Thr Ile
163          50                      55                      60
164      Asn Val Trp Ile Phe Asp Leu Asp Asp Thr Leu Leu Ser Ser Ile Pro
165          65                      70                      75                      80
166      Tyr Tyr Ala Lys Tyr Gly Tyr Gly Thr Glu Asn Thr Ala Ala Gly Ala
167          85                      90                      95
168      Tyr Trp Ser Trp Leu Val Ser Gly Glu Thr Pro Gly Leu Pro Glu Thr
169          100                     105                     110
170      Leu His Leu Tyr Glu Asn Leu Leu Glu Leu Gly Ile Glu Pro Ile Ile
171          115                     120                     125
172      Ile Ser Asp Arg Trp Lys Lys Leu Ser Glu Ile Thr Ile Glu Asn Leu
173          130                     135                     140
174      Lys Ala Val Gly Val Thr Lys Trp Lys His Val Ile Leu Lys Pro Asn
175          145                     150                     155                     160
176      Gly Lys Leu Thr Gln Val Val Tyr Lys Ser Lys Val Arg Asn Gly Leu
177          165                     170                     175
178      Val Arg Gln Gly Tyr Asn Ile Val Gly Ile Ile Gly Asp Gln Trp Ala
179          180                     185                     190
180      Asp Leu Val Glu Asp Thr Pro Gly Arg Val Phe Lys Leu Pro Asn Pro
181          195                     200                     205
182      Leu Tyr Tyr Val Pro Ser
183          210
185 <210> SEQ ID NO: 6
186 <211> LENGTH: 220
187 <212> TYPE: PRT
188 <213> ORGANISM: Arabidopsis thaliana
189 <400> SEQUENCE: 6
190      Ser Ile Asn Tyr Pro Asn Cys Arg Ser Trp His Leu Gly Val Glu Thr
191          1                      5                      10                      15
192      Ser Asn Ile Ile Asn Phe Asp Thr Val Pro Ala Asn Cys Lys Ala Tyr
193          20                      25                      30
194      Val Glu Asp Tyr Leu Ile Thr Ser Lys Gln Tyr Gln Tyr Asp Ser Lys
195          35                      40                      45
196      Thr Val Asn Lys Glu Ala Tyr Phe Tyr Ala Lys Gly Leu Ala Leu Lys
197          50                      55                      60
198      Asn Asp Thr Val Asn Val Trp Ile Phe Asp Leu Asp Asp Thr Leu Leu

```

## RAW SEQUENCE LISTING

DATE: 11/04/2002

PATENT APPLICATION: US/09/478,598

TIME: 16:19:28

Input Set : N:\Crf3\RULE60\09478598.raw

Output Set: N:\CRF4\11042002\I478598.raw

```

199      65      70      75      80
200  Ser Ser Ile Pro Tyr Tyr Ala Lys Tyr Gly Tyr Gly Thr Glu Asn Thr
201      85      90      95
202  Ala Pro Gly Ala Tyr Trp Ser Trp Leu Glu Ser Gly Glu Ser Thr Pro
203      100      105      110
204  Gly Leu Pro Glu Thr Leu Tyr Leu Tyr Glu Asn Leu Leu Glu Leu Gly
205      115      120      125
206  Ile Glu Pro Ile Ile Ile Ser Asp Arg Trp Lys Lys Leu Ser Glu Val
207      130      135      140
208  Thr Val Glu Asn Leu Lys Ala Val Gly Val Thr Lys Trp Lys His Leu
209      145      150      155      160
210  Ile Leu Lys Pro Asn Gly Ser Lys Leu Thr Gln Val Val Tyr Lys Ser
211      165      170      175
212  Lys Val Arg Asn Ser Leu Val Lys Lys Gly Tyr Asn Ile Val Gly Asn
213      180      185      190
214  Ile Gly Asp Gln Trp Ala Asp Leu Val Glu Asp Thr Pro Gly Arg Val
215      195      200      205
216  Phe Lys Leu Pro Asn Pro Leu Tyr Tyr Val Pro Ser
217      210      215      220
219 <210> SEQ ID NO: 7
220 <211> LENGTH: 219
221 <212> TYPE: PRT
222 <213> ORGANISM: Arabidopsis thaliana
223 <400> SEQUENCE: 7
224  Ser Ile Asn Tyr Ala Asn Cys Arg Ser Trp His Leu Gly Val Glu Thr
225      1      5      10      15
226  Ser Asn Ile Ile Asp Phe Asp Thr Val Pro Ala Asn Cys Lys Asp Tyr
227      20      25      30
228  Val Glu Asp Tyr Leu Ile Thr Ser Lys Gln Tyr Gln Tyr Asp Ser Lys
229      35      40      45
230  Thr Val Cys Lys Glu Ala Tyr Phe Tyr Ala Lys Gly Leu Ala Leu Lys
231      50      55      60
232  Asn Asp Thr Val Asn Val Trp Ile Phe Asp Leu Asp Asp Thr Leu Leu
233      65      70      75      80
234  Ser Ser Ile Pro Tyr Tyr Ala Lys Tyr Gly Tyr Gly Thr Glu Lys Thr
235      85      90      95
236  Asp Pro Gly Ala Tyr Trp Leu Trp Leu Gly Thr Gly Ala Ser Thr Pro
237      100      105      110
238  Gly Leu Pro Glu Gly Leu Tyr Leu Tyr Gln Asn Ile Ile Glu Val Gly
239      115      120      125
240  Ile Glu Pro Ile Ile Leu Ser Val Arg Trp Lys Leu Trp Lys Asn Val
241      130      135      140
242  Thr Leu Asn Leu Glu Ala Ala Gly Val Thr Tyr Trp Lys His Leu Ile
243      145      150      155      160
244  Leu Lys Pro Asn Gly Ser Asn Leu Arg Gln Val Val Tyr Lys Ser Lys
245      165      170      175
246  Val Arg Asn Lys Leu Val Lys Lys Gly Tyr Asn Ile Val Gly Asn Ile
247      180      185      190
248  Gly Asp Gln Trp Ala Asp Leu Val Glu Asp Thr Pro Gly Arg Val Phe

```

VERIFICATION SUMMARY

DATE: 11/04/2002

PATENT APPLICATION: US/09/478,598

TIME: 16:19:29

Input Set : N:\Crf3\RULE60\09478598.raw

Output Set: N:\CRF4\11042002\I478598.raw